Towards Model-Driven Sustainability Evaluation

Jörg Kienzle
Gunter Mussbacher
McGill University

Lucy Bastin
Nelly Bencomo
Aston University

Jean-Michel Bruel
Benoît Combemale
University of Toulouse

Christoph Becker
University of Toronto

Stefanie Betz
Furtwangen University

Ruzanna Chitchyan
University of Bristol

Betty H.C. Cheng
Michigan State University

Sonja Klingert
University of Mannheim

Richard F. Paige
McMaster University

Birgit Penzenstadler
Long Beach State University

Norbert Seyff
FHNW & University of Zurich

Eugene Syriani
Université de Montréal

Colin C. Venters
University of Huddersfield

ABSTRACT

Sustainability has emerged as a concern of central relevance. As a wicked problem, it poses challenges to business-as-usual in many areas, including that of modeling. This article addresses a question at the intersection of model-driven engineering and sustainability research: "How can we better support sustainability by bringing together model-driven engineering, data, visualization and self-adaptive systems, to facilitate engagement, exploration, and understanding of the effects that individual and organizational choices have on sustainability?" We explore this question via an idealized vision of an evaluation environment that facilitates integration and mapping of models from multiple diverse sources, visual exploration, and evaluation of what-if scenarios, for stakeholders with divergent perspectives. The article identifies research challenges to be addressed to enable decision making to support sustainability and provides a map of sustainability modeling issues across disciplines.

KEYWORDS

Sustainability, model-driven engineering, model-driven evaluation

1 INTRODUCTION AND BACKGROUND

Sustainability, the capacity to endure [59], has emerged as a concern of central relevance for society. However, the nature of sustainability is distinct from other concerns addressed by computing research, such as automation, self-adaptation or intelligent systems. It demands the consideration of environmental resources, economic prosperity, individual wellbeing, social welfare, and the evolvability of technical systems [15]. Thus, it requires a focus not just on productivity, effectiveness, and efficiency, but also the consideration of longer-term, cumulative, and measurable effects of systemic technology interventions, as well as lateral side-effects not foreseen at the time of implementation. Furthermore, sustainability includes normative elements and encompasses multi-disciplinary aspects and potentially diverging views. As a *wicked problem* (see sidebar A), it challenges business-as-usual in many areas of engineering and computing research.

The complexity of these integrated techno-socio-economic systems and their interactions with the natural environment is driving attention in several areas. These areas include means for understanding the emergent dynamics of these interactions and supporting better decision making through predictive simulation and system adaptation. At the heart of this is the notion of a *model*, an abstraction created for a purpose. Models are used throughout sustainability research (e.g., for hydrology or pollution analysis) as well as software engineering (e.g., for automated code generation). Models have a long history in research related to sustainability. The Global Modeling (GM) initiatives that started in 60's and 70's developed and used large mathematical dynamic global models to simulate large portions of the entire world [22]. GM in general was applied to human decision-making in domains such as economics, policy, defence, minimization of poverty and climate change. The goal of GM is to offer a prediction of the future state of the world, or parts of it, using (perhaps heavily) mathematical equations and assumptions. Mathematical models offer a framework of stability that is useful in domains such as climate modeling, but it may not be the same in the case of social sciences domains.

In GM, several models can be seen as "modules" of a larger one, where outputs of one model are inputs for other(s) model(s). This vision of modularity was perhaps very advanced for its time. The idea of building models of complex systems based on simpler models has progressed enormously in the engineering domains, software engineering included. However, in the areas of social and natural sciences it is not the case [21]. The intention of initiatives related to GM, e.g., International Futures [6] or the GLOBIOM model [5], have common qualities shared by our proposal. However, GM did not present software engineering practices as a relevant aspect, partly due to the state of software engineering in those years.

Model-driven engineering (MDE) [63] advocates the use of models that are successively refined and help analyze system properties. This article addresses a question at the intersection of MDE and sustainability research: "How can we better support and automate sustainability by bringing together models, data, visualization, and self-adaptive systems to facilitate better engagement, exploration, and understanding of the effects that individuals' and organizational choices have on sustainability?". The authors addressed this question with members from the MDE, sustainability design and sustainability modeling communities, building on earlier contributions [27].

The article conducts a focused review of converging research in MDE, data integration, digital curation (see sidebar B), public engagement, and self-adaptive systems with the perspective of sustainability as a driving motivation. We draw upon a vision of a highly capable integrated environment that facilitates integration of models and data from multiple diverse sources and visual exploration of what-if and how-to scenarios for multiple constituencies. This lens is especially effective for such a review due to its central relevance and urgency, but also because of the massively heterogeneous nature of data required to understand sustainability. We note the limitations of existing approaches and the common assumptions around reductionist modeling perspectives, quantification of uncertainty, and resolution of conflicts and contradictions. These issues are leveraged to identify and characterize emerging research challenges.

2 SUSTAINABILITY MODELING

Modeling has been the essential mechanism to cope with complexity. While in science, models are used to describe existing real world phenomena, in engineering, models are mostly used to describe a system that is to be developed. Thus, engineering models are typically constructive, while scientific models, e.g., mathematical models and stochastic models, are typically used to predict real world aspects.

Modeling underpins many activities related to sustainability. As such, research in MDE can provide a framework for conceptualizing and reasoning about sustainability challenges. One key challenge is how to support decision-making and trade-off analysis to guide behavior of (self-adaptive) systems used for addressing sustainability issues. For this purpose, we present an idealized vision of a conceptual model-based framework, termed *Sustainability Evaluation ExperienceR* (SEER, cf. Figure 1). This system enables broader engagement of the community (e.g., scientists, policy makers, general public), facilitates more informed decision-making through what-if scenarios and directly uses these decisions to drive the automatic and dynamic adaptation of self-adaptive systems (SAS) [24]. We elaborate this vision not as a design for a system to be implemented, but as a framework that enables us to distill the main nine capabilities needed to tackle this multidisciplinary challenge. Since we argue that MDE is one of the main enablers for a system like the SEER, we contemplate the challenges for MDE research that lie ahead.

2.1 Vision

This paper introduces the SEER, a conceptual entity that brings together sustainability scientists and decision makers, whose output can be used to guide dynamic adaptation of an SAS. As such, the SEER focuses on 1) enabling scientists to integrate and then test their heterogeneous models with an existing knowledge base; 2) enabling individuals and policy makers to explore economic, social, and environmental impact of decisions, investigate trade-offs and alternatives, and express preferences; 3) automating the acquisition of contextual data and enactment of decisions by directly feeding into the knowledge that guides the adaptation of an SAS. The SEER will give the context to introduce the nine capabilities.

2.1.1 Model Integration.

Scientists need to be able to continuously integrate new knowledge into the SEER in the form of models or data. For example, an agronomist can contribute a biomass growth model corresponding to a newly discovered cultivation technique, or a city can decide to openly disclose urban data. Scientists can further connect this contributed material to available and relevant open data. Furthermore, they could investigate the consistency and validity of their models by testing them in combination with other existing domain models. This would help scientists to reach a common view or to highlight important divergences for discussion. To this end, the SEER must provide facilities for *flexible data and model integration* (C1), *model curation* (C2), as well as *enable trustworthy open-world contributions* (C3). The SEER should also support those scientists in investigating the consistency between heterogeneous models, *accommodating different* and possibly divergent *world views* (C4).

2.1.2 Model Exploration and Investigation.

On the basis of this knowledge, individuals, communities, and policy makers would explore scenarios, evaluate tradeoffs along the five sustainability dimensions [15] (technological, environmental, social, individual, economic) or planetary boundaries [61], and explore direct, enabling and structural effects (see sidebar C). Hence, the SEER must *enable use by the population at large* (C8). For example, a farmer who is considering building a biowaste plant to become energy independent could investigate the consequences of this idea. This analysis needs to include basic information about the farmer's preferences and the current as-is situation, and to elicit any required information which is not available. To analyze this issue, the SEER needs data and model sources, such as an operational model of the farm or the heating system of the house. The SEER visualizes the analysis results to facilitate exploration. For example, economic analysis might suggest that heating with biowaste is more cost effective than oil. However, the user may doubt this assertion and wish to investigate the result, so the SEER should *provide a transparent rationale and quantification of uncertainty* (C6), as well as expose the underlying data. In addition to *generating what-if scenarios* (C5), the SEER should be capable of *generating suggestions* (C7) of how to reach user specified goals including quantifiable impacts.

2.1.3 Model Automation.

Strategic choices typically require a set of well-defined steps to implement them, a process that can benefit significantly from automation. This is especially pertinent when those steps are controlled by an SAS, e.g., smart cities or smart buildings. In such cases, decisions are used directly to drive the run-time adaptation of the SAS. For example, when a farmer chooses to grow a specific crop, the SEER could continuously adjust the irrigation system to deliver the appropriate amount of water to the fields. Thus, the SEER must *perform sustainability evaluation to determine adaptation needs* (C9) to enable broader engagement from the various sustainability stakeholders and would hence serve as an adaptation trigger for an SAS.

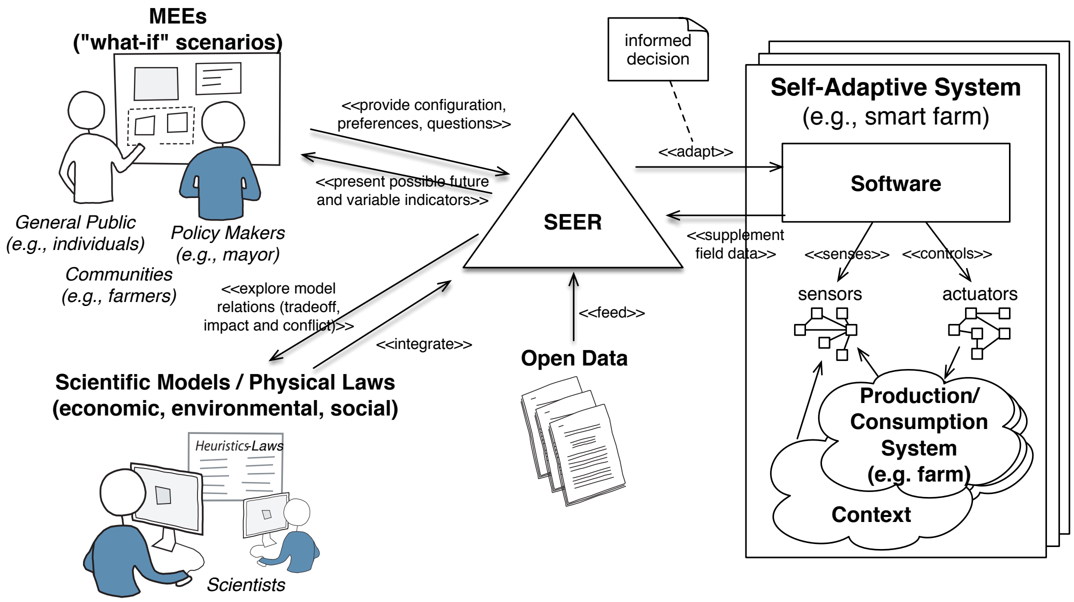

Figure 1: The Sustainability Evaluation ExperienceR

3 MDE FOR SEER

This section revisits the capabilities introduced in the previous section, and discusses how techniques from the MDE community and other associated communities can support them.

MDE aims at raising the level of abstraction at which a software system is designed, implemented and evolved to improve the management of intrinsic complexity [36]. In MDE, a model describes an aspect of a system and is typically created for specific development purposes. Separation of concerns is supported through the use of different modeling languages, each providing constructs based on abstractions that are specific to an aspect of a system. But systems like the SEER also require, as a central function, a set of abilities to curate diverse collections of data and manage them throughout a long-lasting lifecycle to address concerns such as authenticity, archiving, transformation, reuse, appraisal, and preservation. In this context, data monitoring involves the continuous, automated acquisition of new data sets.

3.1 Accommodate Flexible Data and Model Integration (C1)

The MDE community has been investigating how to integrate engineering models for various purposes (e.g., analyses, code generation, simulation, execution). In addition to comparison operators such as those that can be defined in the Epsilon Comparison Language [46], the community has developed various composition operators for model refinement/decomposition [56], model consistency or impact analyses [37] and model merging and weaving [17]. While these composition operators have been extensively studied for homogeneous and structural models [25], recent efforts are also considering behavioral and heterogeneous models [47].

In the software and systems modeling community, research on domain-specific modeling languages (DSMLs) is investigating technologies for developing languages and tools that enable domain experts to develop solutions efficiently. Unfortunately, the current lack of support for explicitly relating concepts expressed in different DSMLs makes it difficult for domain experts to reason about information distributed across models describing different system views. Supporting coordinated use of DSMLs led to the grand challenge of the *globalization of modeling languages* [26] and the GEMOC initiative. Beyond the current investigations that focus on relating languages of similar foundations, sustainability issues will impose additional research challenges relating to multi-scale, uncertainty, and approximation or discontinuity.

An alternative to integrating DSMLs is to integrate models by co-simulation or model translation. For example, the Functional Mock-up Interface (FMI) is a tool independent standard to support both model exchange and co-simulation of dynamic models using a combination of xml-files and compiled C-code [7]. FMI is currently supported by over 130 modeling and simulation tools. Model translation approaches construct model transformation algorithms that integrate models by mapping them into a common modeling formalism. For example the work described in [21] transform System Dynamics models into Discrete Event System Specification (DEVS) models, which can then be integrated further with other discrete modeling formalisms, e.g., state automata.

Unlike software-intensive systems, the SEER requires integration of numerous scientific models, regulations, preferences, etc., when making predictions and in order to consider the many trade-offs when looking for potential solutions. The challenges for integrating models within a SEER are due to the following factors:

Different Foundations: In traditional MDE, foundational notions, e.g., hierarchy/containment or references, are used in constructing models; different notions are used in other modeling spaces (e.g., derived attributes in MetaDepth [32]). The integration process must acknowledge and align these different notions.

Different Technological Spaces: Models may be constructed using mechanisms from different technological spaces (e.g., databases, formulae such as ODEs), with varying assumptions about a) the basic building blocks of modeling, b) how those building blocks can be composed, c) how the well-formedness of models can be established and d) how well-formed models can be manipulated.

Different Levels and Degrees of Abstraction: Integrating models involves more than just establishing a consistent vocabulary: disparate models will use different abstractions (e.g., patterns specific to the type of model), different layers and layering structures (e.g., networking layers versus atmospheric chemistry model layers) and different forms of granularity (e.g., a grid of contemporaneous rainfall observations over a large area versus a time series of measurements of cumulative water flow at one location in a river).

Different Scales: to integrate models at different scales, a model integration approach would have to clearly distinguish:
i) Which models belong to which layers of abstraction (for example, given a predictive model of evapotranspiration that can be constructed for Earth as a whole, for a continent or a watershed, which one is relevant when integrating this with a model of crop production at a country level?)
ii) Which specific model out of a set of alternatives to use when there is no evidence demonstrating superiority of one model over another (for example, with insufficient ground truth to distinguish between two multispectral classifications, what characteristics of the classifiers would help the system to choose an option?)
iii) How conflicts or inconsistencies between models and/or data are resolved (for example, given a set of decision trees which risk being overfitted to their training data, is it necessary to employ an ensemble method such as random forest?)

Different Domains: In order to meaningfully integrate data from a variety of domains, it needs to be carefully described with metadata. This should include descriptions of units, phenomena measured and other conceptual aspects, which are vital for communication when data is released "into the wild".

Composability: A crucial capability for the SEER is to automatically identify which data can and cannot logically be combined. For example, a user might be interested in assessing the economic value of a national park by overlaying its bounds on maps of ecosystem services. Such maps might be calculated in different ways, leading to conflicting results. For example, carbon capture per hectare may be computed for specific land covers by methods which rely on different assumptions about underlying physical processes. Should results derived from such data sets be averaged, or be shown as alternatives? A robust approach to this automated matching requires semantics to describe the underlying worldview implicit in each estimate (see section 3.4).

3.2 Curate and Evolve Models (C2)

The SEER must facilitate continuous management of models to ensure the generation of valid what-if and how-to scenarios. Model management involves supporting updates to models and to model integration. Key activities include model import and creation (e.g., scientific model creation out of data sets), enhancing model quality and representation of different views.

There are two approaches to scientific model creation: either 1) start with a skeletal model with a few initial data points and incrementally collect relevant data while refining the model relationships, or 2) build a model based on analysis of all accessible data.

From the perspective of the robust management of MDE products over time, version control is essential to reflect the state of the model at the time when a data set was imported. When this initial data set does not conform to later, updated versions of the model, maintenance challenges arise for the data sets. Conceptual approaches for version control in MDE have been developed, based on techniques for comparing and differencing models [33] as well as merging models. More recently, tools such as EMF Store [45] and CDO [4] have been developed, which are closely aligned with version control systems such as *git*. Conflicts are common with such approaches, and hence support for detection and resolution are critical. Such tools typically are combined with those for comparison and differencing (for detection), and merging (resolution).

From the perspective of digital curation, larger concerns around provenance, authenticity and stewardship become paramount. The provenance of data has been a central concern in fields such as databases and e-Science [19, 55, 64]. Provenance modelling initiatives have focused on conceptual frameworks for representing generally applicable elements that capture provenance information in standardized ways [69]. Concepts such as *Research Objects* capture more than the data set to support the flexible reuse of various products in research workflows and in particular, model-based scientific workflow software such as Kepler and Taverna [14]. Again, data provenance is a central concern and raises new challenges, as discussed below.

3.3 Enable Trustworthy Open-World Contributions (C3)

To enable trustworthy open-world contributions, everybody should be allowed to contribute to the SEER, regardless of their social background, domains of expertise, or technical qualifications. A simple example of the utility of such contributions are the citizen-science projects. For instance, the UK's *Spring Watch* program enlists radio listeners to report, via text and/or photographs, the observations of native wild life species, which can be a cost-free tool for observing, recording, and where necessary taking actions for preserving biodiversity. Contributions to the SEER would consist not only of data or models, but also of new mappings or relationships for integrating data and models.

To foster trust towards and use of the contributions, their provenance must be publicly availed. This is essential in order to [64]: (a) assure potential data users of the quality of the given data (providing answers to such questions as: what is the data source, were the derivation methods of the current data sound?); (b) support the owners and users with the audit trial (who is using the data? are there any errors in the data generation?); (c) provide recopies for replicating data derivation in order to maintain currency of the data, as well as to maintain clear derivation recipes; (d) support attribution of data for both copyright and liability assignment purposes, and (e) provide information about the data context, and for data discovery.

Currently data curation is being tackled by open-world contributions that have little provenance, so the quality of that data and the collection processes are questionable. For example, in the CARMEN bioinformatics project [1], researchers can submit data and the metadata that describes it. However, provenance information is limited to the identity of the source.Yet, it is widely acknowledged that, in order to provide credible provenance for scientific workflow, one needs to report provenance not only of the provided data (e.g., its sources and their views, including interests, purpose, concepts, principles, knowledge [41]) but also the process through which the data is derived (e.g., used methodology, and technologies for data collection) [41, 64].

In MDE's few open repositories for models, e.g., ReMoDD [35] or the ATL Metamodel Zoo [3], the situation is even worse, as little information is kept on the provenance or quality of the models, despite the long established specification of provenance requirements for e-Science systems [54].

The challenges for trustworthy open-world contributions pertain to:

Subject of Provenance [64], or the provenance of data and its workflow: It is not clear at what level of detail provenance information needs to be gathered (e.g., what granularity should the data be collected, e.g., rainfall per cm^2 or km^2?). Which sources are acceptable, for what purposes [41]? When pulling together several data sets, or starting analysis for a given purpose, are the used data collection methods and technologies compatible / appropriate for the said purpose? Who must take responsibility for errors in data collection or derivation? Eventually, how do the sources, their properties and the workflow affect the data quality, and how can the quality be separated from the notion of provenance itself?

Provenance Representation [64]: Should data be annotated directly with the provenance details (e.g., many scientific workflow tools, such as Taverna, record the provenance data implicitly in event logs [31]), or should provenance be derived at each workflow stage from the previous one? What syntax and semantics should be used to represent it? Can these be applicable across all kinds of domains, as the SEER has to integrate environmental, economic, technical, societal, individual, policy, and cultural aspects of life?

Storing Provenance [64]: What are the costs of collecting and storing the provenance data at various granularity? Clearly, the richer the provenance data, the more it will affect the scalability of data collection and storage.

Integration: If the system accommodates import of new concepts of all kinds, we face integration challenges, for example to find the best, i.e., most reasonable, or most flexible open interfaces and common description language. Furthermore, the research community must let the ontology evolve iteratively, by adding new parts.

Trust: How do you foster trust, or calculate trust into the given model's output? How do we build trust models? How can we apply theoretical research models in the real world while large scale empirical evidence is still missing?

Relationship between Risk and Trust: How to deal with the inherent relationship between risk and trust? What are the risks involved in trusting a given model/data/process, and how to quantify these? Contrary to public perception, high trust does not mean low risk.

Currently research is ongoing on ways for handling many of the above mentioned challenges for *controlled environments*, such as for scientific work flows within tools like Taverna[1] and Kepler[2] (here data sets and workflows are provided only by scientists or models by research groups who stake their professional reputation against the quality of their contributions). When the controls for contributions are removed, however, these challenges redouble and multiply.

[1] https://taverna.incubator.apache.org/

[2] https://kepler-project.org/

3.4 Accommodate Different World Views (C4)

The breadth of the impact of sustainability across five dimensions and multiple time scales, from human to global, inevitably brings with it differing and irreconcilable worldviews, and separates stakeholders socially and temporally.

To avoid bias, the SEER should provide all possible futures accommodating multiple and potentially divergent worldviews to the user given the available data and models. Therefore, the SEER must acknowledge that a model is constructed with its own (often implicit) worldview [51]. Model integration requires combination of the views, which can be challenging or even impossible if they contradict.

The modeling community deals with situations where worldviews are assumed to be consistent across stakeholders if they share the same modelling background [50]. In most engineering environments this is acceptable, since even large-scale systems have an ultimately "bounded" set of stakeholders. In these scenarios, any necessary negotiation of conflicting worldviews is a question of social organization and not addressed in modeling.

Traditional MDE normally resolves contradictions under model integration using constraints and transformations. This is feasible because even when the worldview is not fully shared, there should be overlap arising from agreement on a metamodeling stack (e.g., 3-tiered) and technology (e.g., the Eclipse Modeling Framework (EMF)). This cannot be assumed in modeling for sustainability, where the social structure is so disconnected that the common assumption of consistent worldviews in MDE cannot hold. Different modelling schools need to be integrated and multiple contradictory worldviews need to be made explicit and embraced.

The worldview has to become an explicit part of the modeling infrastructure, and several possible scenarios arise:

Matching Worldviews: In some cases, worldviews can be reconciled. However, there may be no "actual" user/modeler who possesses this integrated view. How can this integrated view be derived/validated?

Incommensurable Worldviews and Models: Considering the fundamentally distinct nature of the types of concerns of interest for the stakeholders in sustainability, perspectives on what seems to be a common concern will not only disagree on the weighting of importance of particular aspects, such as "individual agency", but also on what this concern means, and how to evaluate it.

Contradictory Worldviews: It should not be assumed that reconciliation of contradicting worldviews is always desirable and appropriate. Sometimes it may be desirable and useful to keep track of contradictions between models. To discuss this, we provide a few examples for worldviews that disagree at least partially:

Incommensurable: In California, environmental sustainability can be regarded as fundamentally different in the problem context of preserving existing wetlands versus restoring a urban landscaping back towards the natural desert environment it was taken from.

Contradictory: In many developing cultures, big families still form the heart of the community. In many developed cultures, family structures have been overshadowed by career paths requiring mobility. One consequence is that two-income families struggle with local support systems for their kids while grandparents live far away and struggle with lonely old age. Neither worldview is wrong, but they cannot be consolidated completely.

The research challenge arising from this is not an unrealistic attempt at consolidating all existing worldviews. Instead, what we need are modeling concepts and mechanisms that allow us to contrast different worldviews to illustrate and explore conflicts between the assumptions and implications of two or more worldviews [51]. One option would be to use System Dynamics to reach a group consensus and enhance systems thinking [68] .

However, System Dynamics on its own is arguably incapable of securing consensus [42]. Because it lacks the awareness of social theory required to distinguish consensus from coercion, it must be positioned within a critically aware systems thinking framework that reflects upon its own selectivity, aims to emancipate marginalized perspectives and worldviews, and allows for pluralism in methods and theories [53].

A useful starting direction in tackling these issues could be provided by model documenting guidelines (e.g., the ODD protocol [39, 40]) that help to systematize and disambiguate categorizations of heterogeneous models, though full resolution of integration of such models is an open challenge.

3.5 Generate What-If Scenarios (C5)

The system should support the generation of what-if scenarios based on multiple types of models to project the scenarios' effects with regard to the five sustainability dimensions. Interactive exploration of the scenario as well as the involved data and models should be possible. Here, it is important that the user of the SEER gets a feeling about how a possible future scenario may look and what effects the anticipated changes will have on the different sustainability dimensions. For example, what would a world look like that no longer used fossil fuel? To help SEER users understand the what-if scenarios and make the experience even more tangible, visualization techniques going beyond the presentation of numbers are needed.

What-if scenarios require query formulation, which is supported through query languages. These languages have been investigated by the MDE community with an intensive focus on automatic model management (e.g., constraints, views, transformation). MDE provides languages for expressing structural queries based on first order logic (e.g. OCL [58]), use of optimization and search techniques combined with models [20, 34] as well as for behavioral queries based on temporal logic [52]. These languages rely on the modeling language specification for expressing queries related to the corresponding concepts or their associated behavioral semantics. The concept of Model Experiencing Environments (MEEs) [57] has been introduced as an approach to support complex model and data integration, while offering customizable interfaces to access model analysis results and their visualizations.

The need for broad engagement with diverse communities and decision makers requires an ability to process questions articulated within the mental models and terminologies used by communities, and support cross-domain compatibility and mapping across various domains. Different impacts must be presented back to the user (using different kinds of visualizations), in such a form that the indicators and their underlying assumptions can be deeply and interactively analyzed for a better understanding. Current practices must be adapted to support the what-if scenario capability. This requires a bridging of the gap between the indicators and the modeling concepts manipulated by the SEER. The user must be able to express the indicators of interest, and the specific views to be used for representing them.

3.6 Provide Transparent Reasoning and Quantification of Uncertainty (C6)

If users do not feel that they understand what is happening in a system and why, they are less likely to trust it. Therefore, trustworthiness can only be established if the reasoning provided by the SEER is transparent, meaning that users can understand where data comes from, to what degree it is reliable, and how it is combined in order to generate predictions.

Intra-model relationships have been a general focus of interest in the MDE community. User-defined mappings between MDE models are supported via model management tools such as the Atlas Model Weaver, EMF Compare, or the Epsilon Comparison Language (ECL). These approaches enable users to describe mappings between models and model elements, and attach semantics to the relationships that are produced. Such models are usually within a single technological space (e.g., EMF). There are also software component interface definitions, such as OpenMI and Taverna, which provide APIs that allow models to be configured to exchange data at run-time within workflows. While such technology is meant to be model agnostic, it supports connection of models from within a technological space. Additionally, such frameworks effectively focus on mappings between data, where the models are used to enable the construction of such mappings.

There has been limited research in the MDE community on dynamic model selection from a large set of models or on run-time conflict resolution between models from disconnected domains and disciplines (most conflict resolution has focused on resolution between models from single or related domains). Current work on justifying model integration reasoning is centered around such topics as edit-aware modeling tools that keep track of the steps that the modelers take in modifying the model (e.g., [10]) and tool support that allows one to keep track of all the versions of a model (e.g., Sparx Time Aware Modeling, Magic Draw Comparer, EMFCompare).

In goal modeling, the impact of alternative solutions on stakeholders' objectives is modeled to allow reasoning about trade-offs. Based on such models, explanations may be given of what influences what. It is still challenging to generate clear explanations of scenarios built on top of widely different types of models, each requiring different argumentation and concepts. For example, when analyzing a chart with a Pareto front to make an allocation decision, the farmer might see a cutoff on one dimension. She might ask "but couldn't I do this?", e.g. increase output beyond x? The SEER would need to be able to explain that the Pareto front does not only take into consideration physical possibilities, but also considers legal constraints.

Within the domain of environmental modeling, there has been some consideration of integration challenges, e.g. [71], particularly in relation to the propagation of uncertainty through a series of chained models and its communication in a usable form at the end of the analysis [11]. 'Models' in that context, however complex, are concrete mathematical transformations which represent physical processes such as soil erosion, or non-physical processes such as market fluctuations. As such they are materializations of the more abstract class of models with which the SEER must work, and form just part of the set of components of which it must be composed.

However, many of the insights from the above research also apply to an integrated system such as SEER: for example, the importance of semantics and controlled vocabularies in describing requirements, constraints or phenomena, and the fact that physical models may also be matched and merged as appropriate.

The uncertainty of available data and information hinders the precise specification of certain models and their parameters. Uncertainty may be, for example, epistemic, linguistic or randomized [38] and can derive from many sources including measurement, data transformation, inaccurate definition of the phenomenon of interest or generalizations made to ensure tractable computation. As such, Uncertainty Analysis (UA) and Sensitivity Analysis (SA) are prerequisites for model building [28]. While UA aims to quantify the overall uncertainty associated with the model response as a result of uncertainties in the model input, SA can be used to quantify the impact of parameter uncertainty on the overall simulation/prediction uncertainty. This makes it possible to distinguish between high-leverage variables, whose values have a significant impact on the system behavior, and low-leverage variables, whose values have minimal impact on the system [43, 72]. Such approaches can be used for various purposes, including model validation, evaluating model behavior, estimating model uncertainties, decision-making using uncertain models, and determining potential areas of research [48] and a variety of SA techniques have been developed to achieve such purposes [49]. However, federating several models is likely to result in the potential problem of enlarging the parameter space, which will require the automated detection of hotspots in the parameter space using approaches such as the ones proposed by Danos et al. [30].

Neverthless, not all sources of uncertainty are known, and many are difficult to quantify. Uncertainty which can be assessed statistically may be communicated, for example, using probabilities, which are easily combined across a wide variety of well-supported frameworks and languages, e.g. UncertML [70]. Fuzzy sets are more complex to combine across domains, but can still be represented in mathematical form. However, on many occasions a quality assessment is not easily mapped to a value scale, or a problem does not become apparent until a dataset or model is used, or compared to better alternatives that were not originally available. This is a clearly recognized challenge in citizen science (CS), where a number of initiatives aim to harmonize metadata standards [12], to adapt existing data formats to the CS context [65], to develop robust ontologies to capture heterogeneous data collection protocols and to allow flexible annotation by contributors and expert evaluators alike [9, 13]. Only through such concerted efforts can a potential user assess whether the reliability of a contributed resource matches their criteria, making it fit-for-purpose.

3.7 Generate Suggestions (C7)

The system should be capable of generating suggestions of how to achieve the user's specified goals. This generation of suggestions is based on the capability to create what-if scenarios (C3), as those are needed to build a knowledge base for a recommender system. Based on such a what-if scenario knowledge base, a recommender system can generate how-to scenarios by using model inference. Inferred models can be compared to current ones and criteria applied to select the most appropriate candidate solutions, e.g., the closest to the current situation. Therefore, the SEER has to calculate different alternatives to minimize negative impact on the different sustainability dimensions. To do so, the system has to be informed what a user may and may not change, e.g., they cannot change the weather. Furthermore, the SEER needs to know user preferences in order to make adequate individual suggestions. Such user preferences include the modeling view of the system under consideration, the agency over individual elements, and the scale at which they can be changed. The preferences could even be changed at run-time and the model recalculated based on the updated constraints [66].

3.8 Enable Use by the Population at Large (C8)

Since the SEER is to be used by the population at large, careful consideration must be given to human factors and ergonomics in system design. Some example issues to be addressed here include simple ways to establish and update preferences and goals (e.g., via graphical or voice-based interfaces); results interpretation (e.g., via visualization or voice feedback explaining the results' implications); customization of interactions for different user groups (e.g., domain-specific model customization support for specialist users). The quality of the users' experience [18] should also be considered, accounting for the users' emotional and physiological states, the situational characteristics of the experience and the experience of model use itself [8].

3.9 Evaluating Adaptation for Sustainability (C9)

Based on the sustainability evaluation performed by the SEER, adaptation triggers may be generated to guide the self-adaptation of an SAS. In the original framework proposed by Kephart and Chess [44], an SAS has four key stages (MAPE-K loop): **M**onitoring environment and system conditions, **A**nalysis to determine whether system needs to self-reconfigure, **P**lanning for how to adapt the system safely to satisfy new requirements/needs, and **E**xecution of the adaptation plan. All four stages make use of a **K**nowledge resource. While the original intent for Knowledge was for static information (e.g., sensor properties, policies, constraints, etc.), for our purposes, we realize the Knowledge resource with the SEER. As such, the SEER becomes a dynamic source of sustainability-evaluation knowledge that incorporates input from the stakeholders, scientific models and their integration, open data, results of what-if scenario exploration, user needs, etc. to guide the self-adaptation of an SAS. The entire MAPE-K loop is hence open for human assessment and feedback to derive a recommendation that can either be realized by an automated adaptation or realized by human intervention. For example, Bruel et al. present a smart farming system including an irrigation system that determines and delivers the right amount of water every day in order to maximize produced biomass, based on current water stress, the climate series, biomass models and the farmer's input [16].

4 CONCLUSION

In this article, we explained each capability needed by the SEER in more detail, and reported on how MDE has already contributed towards that capability. However, most of the disciplines in Computer Science (CS) have to come together to realize the SEER vision outlined above. Therefore, we use the ACM Computing Classification System [2] to assess the CS disciplines and create a simplified heat map (see Figure 2) where we indicate for each top-level category whether or not we, i.e., the 16 authors, believe it is not relevant (white), relevant (blue), or highly relevant (red) to realize the SEER. Whenever we feel that some subcategories are notably more important than others, they are mentioned explicitly in the appropriate cells of the heat map. The heat map represents the biased view of the authors, and as a result, the importance of some categories might have been misjudged. In general, it can be supposed that expertise in CS is needed across all capabilities, that each of the CS categories is highly relevant for at least one of the capabilities, and finally that MDE is highly relevant across all capabilities.

REFERENCES

[1] code analysis, repository & modelling for eneuroscience.

[2] Acm computing classification system. https://dl.acm.org/ccs/ccs.cfm, 2018.

[3] The Metamodel Zoo. http://web.emn.fr/x-info/atlanmod/index.php?title=Zoos, last accessed 2019.

[4] CDO Model Repository, last accessed Feb. 2019. http://www.eclipse.org/cdo/.

[5] GLOBIOM. http://www.globiom.org, "Last accessed Feb. 2019".

[6] International Futures. http://pardee.du.edu, "Last accessed Feb. 2019".

[7] The Functional Mockup Interface (FMI) Standard, last accessed Feb. 2019. https://fmi-standard.org.

[8] S. Abrahão et al. User experience for model-driven engineering: Challenges and future directions. In *Model Driven Engineering Languages and Systems*, pages 229–236, 2017.

[9] R. Albertini and A. Isaac. Data on the Web Best Practices: Data Quality Vocabulary. https://www.w3.org/TR/vocab-dqv/.

[10] K. Altmanninger et al. Why Model Versioning Research is Needed!? An Experience Report. In *MoDSE-MCCM Workshop at MoDELS*, pages 1–12, 2009.

[11] L. Bastin et al. Managing uncertainty in integrated environmental modelling: The uncertweb framework. *Environmental Modelling and Software*, 39:116 – 134, 2013. Issue on the Future of Integrated Modeling Science and Technology.

[12] L. Bastin et al. *Good Practices for Data Management*, chapter 11. 2017.

[13] L. Bastin et al. *Volunteered metadata, and metadata on VGI : Challenges and current practices*. 2017.

[14] S. Bechhofer, J. Ainsworth, J. Bhagat, I. Buchan, P. Couch, D. Cruickshank, D. D. Roure, M. Delderfield, I. Dunlop, M. Gamble, C. Goble, D. Michaelides, P. Missier, S. Owen, D. Newman, and S. Sufi. Why Linked Data is Not Enough for Scientists. In *2010 IEEE Sixth International Conference on e-Science*, pages 300–307, Dec. 2010.

[15] C. Becker et al. Requirements: The key to sustainability. *IEEE Software*, 33(1):56–65, Jan 2016.

[16] J.-M. Bruel et al. MDE in Practice for Computational Science. In *Int. Conf. on Computational Science*, June 2015.

[17] G. Brunet et al. A manifesto for model merging. In *Proc. of the 2006 Int. Workshop on Global Integrated Model Management*, GaMMa '06, pages 5–12, 2006.

[18] M. Bui and E. Kemp. E–tail emotion regulation: examining online hedonic product purchases. *Int. J. Retail and Distribution Management*, 41:155–170, 2013.

[19] P. Buneman, S. Khanna, and T. Wang-Chiew. Why and Where: A Characterization of Data Provenance. In *Database Theory ICDT 2001*, LNCS, pages 316–330. Springer, 2001.

[20] C. M. Byers and B. H. Cheng. An approach to mitigating unwanted interactions between search operators in multi-objective optimization. In *Annual Conference on Genetic and Evolutionary Computation*, pages 655–662, 2015.

[21] R. Castro. Open research problems: Systems dynamics, complex systems, chapter 24. In *Theory of Modeling and Simulation (Third Edition)*. Academic Press, 2019.

[22] R. Castro and P. Jacovkis. Computer-Based Global Models: From Early Experiences to Complex Systems. *Journal of Artificial Societies and Social Simulation*, 18(1):1–13, 2015.

[23] P. Checkland. *Systems thinking, systems practice*. 1981.

[24] B. H. C. Cheng et al. Software engineering for self-adaptive systems: A research roadmap. In *Software Engineering for SAS*, pages 1–26, 2009.

[25] M. Clavreul et al. Integrating legacy systems with mde. In *International Conferebce on Software Engineering*, volume 2, pages 69–78, 2010.

[26] B. Combemale et al. Globalizing Modeling Languages. *Computer*, pages 68–71, June 2014.

[27] B. Combemale et al. Modeling for Sustainability. In *Modeling in Software Engineering*, 2016.

[28] M. Crosetto, S. Tarantola, and A. Saltelli. Sensitivity and uncertainty analysis in spatial modelling based on gis. *Agriculture, Ecosystems Environment*, 81(1):71 – 79, 2000.

[29] C. Dallas. Digital curation beyond the wild frontier: a pragmatic approach. *Archival Science*, 16(4):421–457, 2016.

[30] A. Danós, W. Braun, P. Fritzson, A. Pop, H. Scolnik, and R. Castro. Towards an OpenModelica-based Sensitivity Analysis Platform Including Optimization-driven Strategies. In *EOOLT '17*, pages 87–93. ACM, 2017.

[31] S. B. Davidson and J. Freire. Provenance and scientific workflows: Challenges and opportunities. In *International Conference on Management of Data*, SIGMOD '08, pages 1345–1350, New York, NY, USA, 2008. ACM.

[32] J. de Lara and E. Guerra. Deep meta-modelling with metadepth. In *Proc. of the 48th Int. Conference on Objects, Models, Components, Patterns*, TOOLS'10, pages 1–20, 2010.

[33] Dimitrios S. Kolovos and others. Different Models for Model Matching: An analysis of approaches to support model differencing. In *Workshop on Comparison and Versioning of Software Models*, 2009.

[34] M. Faunes et al. Automatically Searching for Metamodel Well-Formedness Rules in Examples and Counter-Examples. In *Model Driven Engineering Languages and Systems*, LNCS, pages 187–202, 2013.

[35] R. France, J. Bieman, and B. H. C. Cheng. Repository for Model Driven Development (ReMoDD). In *Models in Software Engineering*, pages 311–317. Springer, 2007.

[36] R. B. France and B. Rumpe. Model-driven Development of Complex Software: A Research Roadmap. In *Workshop on the Future of Software Engineering (FOSE 2007)*, pages 37–54, 2007.

[37] I. Galvao and A. Goknil. Survey of traceability approaches in model-driven engineering. In *EDOC 2007*, pages 313–313, Oct 2007.

[38] H. Giese et al. *Living with Uncertainty in the Age of Runtime Models*, pages 47–100. 2014.

[39] V. Grimm, U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback. The odd protocol: a review and first update. *Ecological modelling*, 221(23):2760–2768, 2010.

[40] V. Grimm, G. Polhill, and J. Touza. Documenting social simulation models: the ODD protocol as a standard. In *Simulating Social Complexity*, pages 349–365. Springer, 2017.

[41] J. Huang. From big data to knowledge: Issues of provenance, trust, and scientific computing integrity. In *Big Data 2018*, pages 2197–2205, 2018.

[42] M. C. Jackson. *Systems thinking: Creative holism for managers*. Wiley Chichester, 2003.

[43] S. E. Jrgensen and B. D. Fath. 2 - concepts of modelling. In S. E. Jrgensen and B. D. Fath, editors, *Fundamentals of Ecological Modelling*, volume 23 of *Developments in Environmental Modelling*, pages 19 – 93. Elsevier, 2011.

[44] J. O. Kephart and D. M. Chess. The vision of autonomic computing. *Computer*, 36:41–50, Jan 2003.

[45] M. Koegel and J. Helming. EMFStore: A Model Repository for EMF Models. In *ICSE 2010 - Volume 2*, ICSE '10, pages 307–308. ACM, 2010.

[46] D. S. Kolovos. Establishing correspondences between models with the epsilon comparison language. In *Model Driven Architecture - Foundations and Applications*, pages 146–157, 2009.

[47] V. Larsen et al. A Behavioral Coordination Operator Language (BCOoL). In *MODELS 2015*, Aug. 2015.

[48] W. Lehr, D. Calhoun, R. Jones, A. Lewandowski, and R. Overstreet. Model sensitivity analysis in environmental emergency management: a case study in oil spill modeling. In *Winter Simulation Conference*, pages 1198–1205, Dec 1994.

[49] D. M Hamby. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental monitoring and assessment*, 32:135–154, 09 1994.

[50] D. Meadows, J. Richardson, and G. Bruckmann. *Groping in the Dark: The First Decade of Global Modelling*. John Wiley & Sons, 1982.

[51] D. H. Meadows and J. M. Robinson. The electronic oracle: computer models and social decisions. *System Dynamics Review*, 18(2):271–308.

[52] B. Meyers et al. Promobox: A framework for generating domain-specific property languages. In *Software Language Engineering*, pages 1–20, 2014.

[53] G. Midgley. What Is This Thing Called CST? In *Critical Systems Thinking*, pages 11–24. Springer, Boston, MA, 1996.

[54] S. Miles, P. Groth, M. Branco, and L. Moreau. The requirements of using provenance in e-science experiments. *Journal of Grid Comp.*, 5(1):1–25, 2007.

[55] L. Moreau, P. Groth, S. Miles, J. Vazquez-Salceda, J. Ibbotson, S. Jiang, S. Munroe, O. Rana, A. Schreiber, V. Tan, and L. Varga. The Provenance of Electronic Data. *Commun. ACM*, 51(4):52–58, Apr. 2008.

[56] G. Mussbacher et al. Assessing composition in modeling approaches. In *CMA Workshop*, CMA '12, pages 1:1–1:26, 2012.
[57] G. Mussbacher et al. The relevance of model-driven engineering thirty years from now. In *MODELS 2014*, volume 8767 of *LNCS*, pages 183–200, 2014.
[58] Object Management Group (OMG). *Object Constraint Language (v2.4)*, 2014.
[59] OED Online. Oxford English Dictionary Online. http://www.oed.com/view/Entry/299890, Oct. 2017.
[60] H. W. Rittel and M. M. Webber. Dilemmas in a General Theory of Planning. *Policy Sciences*, 4(2):155–169, 1973.
[61] J. Rockström et al. A safe operating space for humanity. *Nature*, 461:472–475, Sep 2009.
[62] C. Rusbridge et al. The Digital Curation Centre: A Vision for Digital Curation. In *International Symposium on Mass Storage Systems and Technology*, LGDI '05, pages 31–41, 2005.
[63] D. C. Schmidt. Model-driven engineering. *IEEE Computer*, 39(2):25–31, 2006.
[64] Y. L. Simmhan, B. Plale, and D. Gannon. A survey of data provenance in e-science. *SIGMOD Rec.*, 34(3):31–36, Sept. 2005.
[65] I. Simonis et al. Sensor Web Enablement (SWE) for Citizen Science. In *Proc. of the IEEE Int. Geoscience and Remote Sensing Symposium*, 2016.
[66] U. Tikhonova et al. Constraint-based Run-time State Migration for Live Modeling. In *Software Language Engineering*, 2018.
[67] W. Ulrich. *Critical Heuristics of Social Planning : A New Approach to Practical Philosophy*. Number 3 in Schriftenreihe des Management. Bern, 1983.
[68] J. A. M. Vennix. Building consensus in strategic decision making: System dynamics as a group support system. *Group Decision and Negotiation*, 4(4):335–355, Jul 1995.
[69] W3C Working Group. PROV-Overview.
[70] M. Williams et al. Uncertml: An xml schema for exchanging uncertainty. In *Proc. of the 16th Conference GISRUK 2008*, pages 275–279, 2008.
[71] D. Wirtz and W. Nowak. The rocky road to extended simulation frameworks covering uncertainty, inversion, optimization and control. *Environmental Modelling and Software*, 93:180–192, 2017.
[72] Y. Zheng, F. Han, Y. Tian, B. Wu, and Z. Lin. Chapter 5 - addressing the uncertainty in modeling watershed nonpoint source pollution. In *Ecological Modelling and Engineering of Lakes and Wetlands*, volume 26 of *Developments in Environmental Modelling*, pages 113 – 159. Elsevier, 2014.

A WICKED PROBLEMS

The concept of 'Wicked problems' was first described in the context of planning by Rittel and Webber [60]. The concept has been used in sustainability-related domains to conceptualize issues including climate change, controlling pandemics and reducing social injustice. Often misunderstood within computing as simply 'difficult' problems, the concept rather points to the inadequacy of problem-solution pairs when used to identify and address complex issues in such situations. The crucial challenge in many situations lies instead in the multiplicity of legitimate and legitimately contradictory perspectives and worldviews about what the issues are. Those views cannot simply be reduced to a 'correct' problem definition using logical operations, but require a discursive process to articulate a definition of the issues to address [23, 60, 67].

B DIGITAL CURATION

Data curation is a type of digital curation, which involves the active management and preservation of digital resources for future use . Digital resources extend far beyond what is commonly understood as data and include artifacts as diverse as scientific models, engineering models, and electronic records. Curation aims to ensure the quality of resources and provide a record of provenance to make resources discoverable and meaningful and instill trust in their authenticity. The ability to effectively create, share and manage diverse assets for current and future use is critical for a sustainable society. Supporting trust provides a crucial objective for curation activities, but curation applies not only to resources that are assumed to be trusted a priori. As Rusbridge et al. point out, 'long term stewardship of digital assets is the responsibility of everyone in the digital information value chain' [62]. Crucial curation activities may be carried out by actors that are not information professionals, such as citizen scientists annotating and releasing a data set. This is especially pertinent in our scenario, where many activities of curating data sets and models take place "in the wild" [29] beyond the narrowly controlled confines of a rigorously defined data curation workflow.

C ORDERS OF EFFECTS ON SUSTAINABILITY

Any given (software) system exercises three types of effects on sustainability of its situated environment [15]: *Immediate effect* occurs due to the production and immediate use of the system, e.g., direct environmental impact of the SEER includes the amount of energy and effort spent on the development, and the reduction of energy consumed by using SEER to find (and set up) a low energy boiler.

Enabling effect comes from the ongoing use of a system, e.g., as SEER promotes more energy efficient choices to all its users, the system would enable reduction of energy use for heating, lighting, transportation, etc., which cuts down the amount of energy resources needed. Yet, it also increases use of broadband for model evaluation, possibly require additional servers to process the vast number of models that the open participation of users from several domains requires, thus increasing energy and broadband consumption.

Structural effects arise due to the long-term reaction of the dynamic socio-economic system to the presence and use of the system, including lifestyle and economic/structural changes. For instance, if many farmers look for an energy efficient way of selling their produce, with SEER recommending e-trade, the trade may move from physical markets to e-shops and e-markets, thus changing the selling and shopping behaviors as well as markets in a given community.

Capabilities

ACM Computing Classification System	**C1 (Enable Flexible Model Integration and Monitoring)**	**C2 (Model Curation)**	**C3 (Enable Trustworthy Open-World Contributions)**	**C4 (Accommodate Different World Views)**	**C5 (Generation of what-if scenarios)**	**C6 (Transparent Reasoning / Uncertainty)**	**C7 (Generate suggestions)**	**C8 (Accessible to the population at large)**	**C9 (Sustainability Evaluation-based Adaptation)**
Hardware		Storage	Hardware Validation			Robustness			
Computer Systems Organization	Distributed Systems, Real-Time, CPS		Distributed Systems, Real-Time			Real-Time, Dependable Systems			Distributed Systems, Real-Time, CPS
Networks		Network Reliability				Network Reliability			
Software and Its Engineering (Languages, ...)	MDE	MDE	MDE, Formal Methods	MDE, Formal Methods	MDE	MDE	MDE	MDE, Context-Specific Languages	MDE
Theory of Computation		Database Theory	Graph Algorithms Analysis	Timed and Hybrid Models, Database Theory		Approximation Algorithms Analysis, Separation Logic			
Mathematics of Computing									
Information Systems	Data Management Systems, Spatial-Temporal Systems	Data Management Systems, Storage Management		Data Management Systems (DMS)	DMS, Data Mining, Decision Support Systems	Data Management Systems	Data Management Systems	Users and Interactive Retrieval	
Security and Privacy	Sec. in Hardware, Systems Sec., Network Sec., Software and Application Sec.	Database and Storage Security	Security Services, Intrusion/anomaly Detection			Security Services, Intrusion/anomaly Detection		Software and Application Sec., Human and Societal Aspects of Sec. and Privacy	Security in Hardware
Human-centered Computing	Interaction Design, Visualization			Interaction Design, Visualization		Interaction Design, Visualization			
Computing Methodologies	AI, ML, Mod. and Sim., Dist. Computing Methodologies			Modeling and Simulation	Modeling and Simulation	Modeling and Simulation	Modeling and Simulation	Computer Graphics, Modeling and Simulation	Distributed Computing Methodologies
Applied Computing									
Social and Professional Topics			Computing / Technology Policy	Computing / Technology Policy, User Characteristics	Computing and Business		Computing and Business	Comp. Educ., Comp. / Techn. Policy, User Characteristics	